

SEQUENCE LISTING

(i) APPLICANT: Lebrun, Michel
Sailland, Alain
Freyssinet, Georges
DeGruae, Eric

(ii) TITLE OF INVENTION: Mutated 5-enolpyruvylshikimate-3-phosphate synthase, gene coding for this protein and transformed plants containing this gene

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Connolly, Bove, Lodge and Hutzel
(B) STREET: 1220 Market Street
(C) CITY: Wilmington
(D) STATE: Delaware
(E) COUNTRY: U.S.A.
(F) ZIP: 19899

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE:
(B) COMPUTER:
(C) OPERATING SYSTEM:
(D) SOFTWARE:

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/FR96/01128
(B) FILING DATE: 18-JUL-1996

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Robert G. McMorro, Jr.

(viii) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (302) 658-9141
(B) TELEFAX: (302) 658-5614

(2) INFORMATION FOR SEQ ID NO:1:

11) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1713 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA

(v1) ORIGINAL SOURCE:
(A) ORGANISM: Zoo says
(B) STRAIN: Black Mexican Sweet
(F) TISSUE TYPE: Callus

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: lambda gt10
(B) CLONE: pRPA-ML-711

(K1) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATCAATTTC	ACACAGAGAA	CAGCTATGAC	CATGATTACG	AATTCGGGOC	CGGCGCGCTG	60
ATCGCGCGCG	GGCAGCGCGC	GCGCGCGCTC	AGCGCGCTC	CGAGGAGATC	GTGCTGCACG	120
CCATCAAGGA	GATCTCGCGC	ACCGTCAAGC	TGCGCGGCTC	CAAGTGGCTT	TCCACCGGA	180
TGCTCTACT	CGCGCGCTG	TGCGAGGGA	CAACAGTGT	TGATAAGCTG	CTGAACAGTG	240
AGGATGTCCA	CTACATGCTC	GCGCGCTTGA	GGACTCTTGG	TCTCTCTGTC	GAAGCGGACA	300
AAGCTGCCA	AAGAGCTGTA	GTGTGTGGCT	GTGTGTGAAA	GTTCGCAGTT	GAGGAAGCTA	360
AGAGGGAAT	CGAGCTCTTC	TGCGCGAATG	CTGAACTGTC	AATCGCGCCA	TTCACAGCAG	420
CTGTACTGTC	TGCTGTGGA	AATGCACTT	ACGTGCTTGA	TGGAATACCA	AGAATGAGGG	480

AAAAAAAT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 540
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 560
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 580
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 600
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 620
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 640
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 660
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 680
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 700
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 720
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 740
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 760
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 780
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 800
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 820
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 840
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 860
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 880
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 900
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 920
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 940
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 960
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 980
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1000
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1020
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1040
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1060
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1080
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1100
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1120
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1140
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1160
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1180
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1200
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1220
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1240
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1260
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1280
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1300
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1320
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1340
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1360
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1380
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1400
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1420
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1440
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1460
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1480
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1500
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1520
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1540
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1560
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1580
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1600
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1620
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1640
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1660
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1680
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1700
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1720
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1740
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1760
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1780
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1800
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1820
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1840
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1860
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1880
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1900
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1920
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1940
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1960
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 1980
 TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT TTTTATTTT 2000

(2) INFORMATION FOR SEQ ID NO:2:

- (1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1340 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(11) MOLECULE TYPE: cDNA

(v1) ORIGINAL SOURCE:
 (A) ORGANISM: *Zea mays*
 (B) STRAIN: Black Mexican Sweet

(v11) IMMEDIATE SOURCE:
 (B) CLONE: pRPA-ML-716

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 6..1337

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCATG GCC GGC GCC GAG GAG ATC GTG CTG CAG CCC ATC AAG GAG ATC 47
 Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile
 1 5 10
 TCC GGC ACC GTC AAG CTG CCG GGG TCC AAG TCC CTT TCC AAC CCG ATC 95
 Ser Gly Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile
 15 20 25 30
 CTC CTA CTC GCC GGC CTG TCC GAG GGG ACA ACA GTG GTT GAT AAC CTG 143
 Leu Leu Leu Ala Ala Leu Ser Glu Gly Thr Thr Val Val Asp Asn Leu
 35 40 45
 CTG AAC AGT GAG GAT GTC CAC TAC ATG CTC GGG GGC TTG AGG ACT CTT 191
 Leu Asn Ser Glu Asp Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu

40	55	60	
GTT CTC TCT TTC GAA GGT GAC AAA GCT GCC AAA AGA GCT GTA GTT GTT Gly Leu Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val 65 70 75	239		
GGT TTT GGT GGA AAG TTC CCA GTT GAG GAT GCT AAA GAG GAA GTG CAG Gly Cys Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln 80 85 90	287		
CTC TTC TTG GGS AAT GCT GGA ACT GCA ATG CCG CCA TTG ACA GCA GCT Leu Phe Leu Gly Asn Ala Gly Thr Ala Met Arg Pro Leu Thr Ala Ala 95 100 105 110	335		
GTT ACT GCT GCT GGT GGA AAT GCA ACT TAC GTG CTT GAT GGA GTA CCA Val Thr Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro 115 120 125	383		
AGA ATG AGG GAG AGA CCG ATT GGC GAC TTG GTT GTC GGA TTG AAG CAG Arg Met Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln 130 135 140	431		
CTT GGT GCA GAT GTT GAT TGT TTC CTT GGC ACT GAC TGC CCA CCT GTT Leu Gly Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val 145 150 155	479		
GGT GTC AAT GGA ATC GGA GGG CTA CTT GGT GGC AAG GTC AAG CTG TCT Arg Val Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser 160 165 170	527		
GGC TCC ATC AGC AGT CAG TAC TTG AGT GGC TTG CTG ATG GCT GCT CCT Gly Ser Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro 175 180 185 190	575		
TTG GCT CTT GGG GAT GTG GAG ATT GAA ATC ATT GAT AAA TTA ATC TCC Leu Ala Leu Gly Asp Val Glu Ile Glu Ile Ile Asp Lys Leu Ile Ser 195 200 205	623		
ATT CCG TAC GTC GAA ATG ACA TTG AGA TTG ATG GAG CTT TTT GGT GTG Ile Pro Tyr Val Glu Met Thr Leu Arg Leu Met Glu Arg Phe Gly Val 210 215 220	671		
AAA GCA GAG CAT TCT GAT AGC TGG GAC AGA TTC TAC ATT AAG GGA GGT Lys Ala Glu His Ser Asp Ser Trp Asp Arg Phe Tyr Ile Lys Gly Gly 225 230 235	719		
CAA AAA TAC AAG TCC CCT AAA AAT GGC TAT GTT GAA GGT GAT GGC TCA Gln Lys Tyr Lys Ser Pro Lys Asn Ala Tyr Val Glu Gly Asp Ala Ser 240 245 250	767		
AGC GCA AGC TAT TTC TTG GCT GGT GCT GCA ATT ACT GGA GGG ACT GTG Ser Ala Ser Tyr Phe Leu Ala Gly Ala Ala Ile Thr Gly Gly Thr Val 255 260 265 270	815		
ACT GTG GAA GGT TGT GGC ACC ACC AGT TTG CAG GGT GAT GTG AAG TTT Thr Val Glu Gly Cys Gly Thr Thr Ser Leu Gln Gly Asp Val Lys Phe 275 280 285	863		
GCT GAG GTA CTG GAG ATG ATG GGA GCG AAG GTT ACA TGG ACC GAG ACT Ala Glu Val Leu Glu Met Met Gly Ala Lys Val Thr Trp Thr Glu Thr 290 295 300	911		
AGC GTA ACT GTT ACT GGC CCA CCG CCG GAG CCA TTT GGG AGG AAA CAC Ser Val Thr Val Thr Gly Pro Pro Arg Glu Pro Phe Gly Arg Lys His 305 310 315	959		
CTC AAG GCG ATT GAT GTC AAC ATG AAC AAG ATG CCT GAT GTC GCC ATG Leu Lys Ala Ile Asp Val Asn Met Asn Lys Met Pro Asp Val Ala Met 320 325 330	1007		
ACT CTT GCT GTG GTT GGC CTC TTT GGC GAT GGC CCG ACA GGC ATC AGA Thr Leu Ala Val Val Ala Leu Phe Ala Asp Gly Pro Thr Ala Ile Arg 335 340 345 350	1055		
GAC GTG GCT TCC TGG AGA GTA AAG GAG AOC GAG AGG ATG GTT GCG ATC Asp Val Ala Ser Trp Arg Val Lys Glu Thr Glu Arg Met Val Ala Ile 355 360 365	1103		
CCG ACC GAG CTA ACC AAG CTG GGA GCA TGT GTT GAG GAA GCG CCG GAC Arg Thr Glu Leu Thr Lys Leu Gly Ala Ser Val Glu Glu Gly Pro Asp 370 375 380	1151		

380

TAA 1340

(X1) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile Ser Gly
 1 5 10 15
 Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile Leu Leu
 20 25 30
 Leu Ala Ala Leu Ser Glu Gly Thr Thr Val Val Asp Asn Leu Leu Asn
 35 40 45
 Ser Glu Asp Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu Gly Leu
 50 55 60
 Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val Gly Cys
 65 70 75 80
 Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln Leu Phe
 85 90 95
 Leu Gly Asn Ala Gly Thr Ala Met Arg Pro Leu Thr Ala Ala Val Thr
 100 105 110
 Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro Arg Met
 115 120 125
 Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln Leu Gly
 130 135 140
 Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val Arg Val
 145 150 155 160
 Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser Gly Ser
 165 170 175
 Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro Leu Ala
 180 185 190
 Leu Gly Asp Val Glu Ile Glu Ile Ile Asp Lys Leu Ile Ser Ile Pro
 195 200 205
 Tyr Val Glu Met Thr Leu Arg Leu Met Glu Arg Phe Gly Val Lys Ala
 210 215 220
 Glu His Ser Asp Ser Trp Asp Arg Phe Tyr Ile Lys Gly Gly Gln Lys
 225 230 235 240
 Tyr Lys Ser Pro Lys Asn Ala Tyr Val Glu Gly Asp Ala Ser Ser Ala
 245 250 255

(2) INFORMATION FOR SEQ ID NO:4:

(A) LENGTH: 1340 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(v1) ORIGINAL SOURCE:

(A) ORGANISM: See page
(B) STRAIN: Black Mexican Sweet

(B) CLONE: PRPA-ML-720

(A) NAME/KEY: CDS
(B) LOCATION: 6..1337

CCATG	GCC	GCC	GAG	GAG	ATC	GTG	CTG	CAG	CCC	ATC	AAG	GAG	ATC		47	
Ala	Gly	Ala	Glu	Glu	Ile	Val	Leu	Gln	Pro	Ile	Lys	Glu	Ile			
1					5				10							
TCC	GCC	ACC	GTC	AAG	GTG	GGG	GGA	TCC	AAG	TGG	CTT	TCC	AAC	CGG	ATC	95
Ser	Gly	Thr	Val	Lys	Leu	Pro	Gly	Ser	Lys	Ser	Leu	Ser	Asn	Arg	Ile	
15					20				25					30		
CTC	CTA	CTC	GCC	GCC	CTG	TCC	GAG	GGG	ACA	GTG	GTT	GAT	AAC	CTG		163
Leu	Leu	Leu	Ala	Ala	Leu	Ser	Glu	Gly	Thr	Thr	Val	Val	Asp	Asn	Leu	
				35				40					45			
CTG	AAC	AGT	GAG	GAT	GTC	CAC	TAC	ATG	CTC	GGG	GCC	TGG	AGG	ACT	CTT	181
Leu	Asn	Ser	Glu	Asp	Val	His	Tyr	Met	Leu	Gly	Ala	Leu	Arg	Thr	Leu	
			50					55					60			
GGT	CTC	TCT	GTC	GAA	GCG	GAC	AAA	GCT	GCC	AAA	AGA	GCT	GTA	GTT	GTT	239
Gly	Leu	Ser	Val	Glu	Ala	Asp	Lys	Ala	Ala	Lys	Arg	Ala	Val	Val	Val	
	65						70					75				

GAC	TTT	GCT	GGA	AAG	TTC	GCA	GTT	GAG	GAT	GCT	AAA	GAG	GAA	GTC	GAG	287
Gly	Cys	Gly	Gly	Lys	Phe	Pro	Val	Glu	Asp	Ala	Lys	Glu	Glu	Val	Gln	
40						45					90					
CTC	TTG	TTG	GGG	AAT	GCT	GGA	ATC	GCA	ATG	GAG	TTC	TTG	ACA	GCA	GCT	335
Leu	Phe	Leu	Gly	Asn	Ala	Gly	Ile	Ala	Met	Arg	Ser	Leu	Thr	Ala	Ala	
55					100					105					110	
GTT	ACT	GCT	GCT	GGT	GGA	AAT	GCA	ACT	TAC	GTG	CTT	GAT	GGA	GTA	CCA	383
Val	Thr	Ala	Ala	Gly	Gly	Asn	Ala	Thr	Tyr	Val	Leu	Asp	Gly	Val	Pro	
				115					120					125		
AGA	ATG	AGG	GAG	AGA	CCC	ATT	GGC	GAC	TTG	GTT	GTG	GGA	TTG	AAG	CAG	431
Arg	Met	Arg	Glu	Arg	Pro	Ile	Gly	Asp	Leu	Val	Val	Gly	Leu	Lys	Gln	
			130					135					140			
CTT	GGT	GCA	GAT	GTT	GAT	TGT	TTC	CTT	GGC	ACT	GAC	TGC	CCA	CCT	GTT	479
Leu	Gly	Ala	Asp	Val	Asp	Cys	Phe	Leu	Gly	Thr	Asp	Cys	Pro	Pro	Val	
			145				150					155				
GCT	GTG	AAT	GGA	ATC	GGA	GGG	CTA	GCT	GGT	GGC	AAG	GTG	AAG	CTG	TCT	527
Arg	Val	Asn	Gly	Ile	Gly	Gly	Leu	Pro	Gly	Gly	Lys	Val	Lys	Leu	Ser	
	160				165						170					
GGC	TCC	ATC	AGC	AAG	CAG	TAC	TTG	AAT	GCC	TTG	CTG	ATG	GCT	GCT	CCT	575
Gly	Ser	Ile	Ser	Ser	Gln	Tyr	Leu	Ser	Ala	Leu	Leu	Met	Ala	Ala	Pro	
175					180					185					190	
TTG	GCT	CTT	GGG	GAT	GAT	GAG	ATT	GAA	ATC	ATT	GAT	AAA	TTA	ATC	TCC	623
Leu	Ala	Leu	Gly	Asp	Val	Glu	Ile	Glu	Ile	Ile	Asp	Lys	Leu	Ile	Ser	
			195						200					205		
ATT	CCG	TAC	GTG	GAA	ATG	ACA	TTG	AGA	TTG	ATG	GAG	GCT	TTT	GCT	GTG	671
Ile	Pro	Tyr	Val	Glu	Met	Thr	Leu	Arg	Leu	Met	Glu	Arg	Phe	Gly	Val	
		210					215						220			
AAA	GCA	GAG	CAT	TCT	GAT	AAC	TGG	GAC	AGA	TTC	TAC	ATT	AAG	GGA	GGT	719
Lys	Ala	Glu	His	Ser	Asp	Ser	Trp	Asp	Arg	Phe	Tyr	Ile	Lys	Gly	Gly	
		225					230					235				
CAA	AAA	TAC	AAG	TCC	CCT	AAA	AAT	GCC	TAT	GTT	GAA	GGT	GAT	GCC	TCA	767
Gln	Lys	Tyr	Lys	Ser	Pro	Lys	Asn	Ala	Tyr	Val	Glu	Gly	Asp	Ala	Ser	
	240				245					250						
AGC	GCA	AGC	TAT	TTG	TTG	GCT	GCT	GCT	GCA	ATT	ACT	GGA	GGG	ACT	GTG	815
Ser	Ala	Ser	Tyr	Phe	Leu	Ala	Gly	Ala	Ala	Ile	Thr	Gly	Gly	Thr	Val	
255					260					265					270	
ACT	GTG	GAA	GCT	TGT	GGC	ACC	ACC	AAT	TTG	CAG	GGT	GAT	GTG	AAG	TTT	863
Thr	Val	Glu	Gly	Cys	Gly	Thr	Thr	Ser	Leu	Gln	Gly	Asp	Val	Lys	Phe	
			275						280					285		
GCT	GAG	GTA	CTG	GAG	ATG	ATG	GGA	GGG	AAG	GTT	ACA	TGG	ACC	GAG	ACT	911
Ala	Glu	Val	Leu	Glu	Met	Met	Gly	Ala	Lys	Val	Thr	Trp	Thr	Glu	Thr	
		290					295						300			
AGC	GTA	ACT	GTT	ACT	GGC	CCA	CCG	CCG	GAG	CCA	TTT	GGG	AGG	AAA	CAC	959
Ser	Val	Thr	Val	Thr	Gly	Pro	Pro	Arg	Glu	Pro	Phe	Gly	Arg	Lys	His	
		305					310					315				
CTC	AAG	GCG	ATT	GAT	GTG	AAC	ATG	AAC	AAG	ATG	CCT	GAT	GTG	GCC	ATG	1007
Leu	Lys	Glu	Ile	Asp	Val	Asn	Met	Asn	Lys	Met	Pro	Asp	Val	Ala	Met	
	320					325					330					
ACT	GTT	GCT	GTG	GTT	GCC	CTC	TTT	GCC	GAT	GGC	CCG	ACA	GCC	ATC	AGA	1055
Thr	Leu	Ala	Val	Val	Ala	Leu	Phe	Ala	Asp	Gly	Pro	Thr	Ala	Ile	Arg	
					340				345					350		
GAC	GTG	GCT	TCC	TGG	AGA	GTA	AAG	GAG	ACC	GAG	AAG	ATG	GTT	GCG	ATC	1103
Asp	Val	Ala	Ser	Trp	Arg	Val	Lys	Glu	Thr	Glu	Arg	Met	Val	Ala	Ile	
			355					360						365		
CCG	ACG	GAG	CTA	ACC	AAG	CTG	GGA	GCA	TCT	GTT	GAG	GAA	GGG	CCG	GAC	1151
Arg	Thr	Glu	Leu	Thr	Lys	Leu	Gly	Ala	Ser	Val	Glu	Glu	Gly	Pro	Asp	
			370				375						380			
TAC	TGC	ATC	ATC	ACG	CCG	CCG	GAG	AAG	CTG	AAC	GTG	ACG	CCG	ATC	GAC	1199
Tyr	Cys	Ile	Ile	Thr	Pro	Pro	Glu	Lys	Leu	Asn	Val	Thr	Ala	Ile	Asp	
							390						395			

(2) INFORMATION FOR SEQ ID NO:5:

(A) LENGTH: 444 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala Gly Ala Glu Glu Ile Val Leu Gln Pro Ile Lys Glu Ile Ser Gly
 1 5 10 15
 Thr Val Lys Leu Pro Gly Ser Lys Ser Leu Ser Asn Arg Ile Leu Leu
 20 25 30
 Leu Ala Ala Leu Ser Glu Gly Thr Thr Val Val Asp Asn Leu Leu Asn
 35 40 45
 Ser Glu Asp Val His Tyr Met Leu Gly Ala Leu Arg Thr Leu Gly Leu
 50 55 60
 Ser Val Glu Ala Asp Lys Ala Ala Lys Arg Ala Val Val Val Gly Cys
 65 70 75 80
 Gly Gly Lys Phe Pro Val Glu Asp Ala Lys Glu Glu Val Gln Leu Phe
 85 90 95
 Leu Gly Asn Ala Gly Ile Ala Met Arg Ser Leu Thr Ala Ala Val Thr
 100 105 110
 Ala Ala Gly Gly Asn Ala Thr Tyr Val Leu Asp Gly Val Pro Arg Met
 115 120 125
 Arg Glu Arg Pro Ile Gly Asp Leu Val Val Gly Leu Lys Gln Leu Gly
 130 135 140
 Ala Asp Val Asp Cys Phe Leu Gly Thr Asp Cys Pro Pro Val Arg Val
 145 150 155 160
 Asn Gly Ile Gly Gly Leu Pro Gly Gly Lys Val Lys Leu Ser Gly Ser
 165 170 175
 Ile Ser Ser Gln Tyr Leu Ser Ala Leu Leu Met Ala Ala Pro Leu Ala
 180 185 190
 Leu Gly Asp Val Glu Ile Glu Ile Ile Asp Lys Leu Ile Ser Ile Pro
 195 200 205
 Tyr Val Glu Met Thr Leu Arg Leu Met Glu Arg Phe Gly Val Lys Ala
 210 215 220
 Glu His Ser Asp Ser Trp Asp Arg Phe Tyr Ile Lys Gly Gly Gln Lys
 225 230 235 240
 Tyr Lys Ser Pro Lys Asn Ala Tyr Val Glu Gly Asp Ala Ser Ser Ala
 245 250 255
 Ser Tyr Phe Leu Ala Gly Ala Ala Ile Thr Gly Gly Thr Val Thr Val
 260 265 270
 Glu Gly Cys Gly Thr Thr Ser Leu Gln Gly Asp Val Lys Phe Ala Glu
 275 280 285

Val	Leu	Glu	Met	Met	Gly	Ala	Lys	Val	Thr	Thr	Thr	Ser	Val	
210						295						300		
Thr	Val	Thr	Gly	Pro	Pro	Arg	Glu	Pro	Phe	Gly	Arg	Lys	His	Leu
304					310					315				Lys
														320
Ala	Ile	Asp	Val	Asn	Met	Asn	Lys	Met	Pro	Asp	Val	Ala	Met	Thr
				325					330					Leu
														335
Ala	Val	Val	Ala	Leu	Phe	Ala	Asp	Gly	Pro	Thr	Ala	Ile	Arg	Asp
			340					345					350	Val
Ala	Ser	Trp	Arg	Val	Lys	Glu	Thr	Glu	Arg	Met	Val	Ala	Ile	Arg
		355					360					365		Thr
Glu	Leu	Thr	Lys	Leu	Gly	Ala	Ser	Val	Glu	Glu	Gly	Pro	Asp	Tyr
	370					375					380			Cys
Ile	Ile	Thr	Pro	Pro	Glu	Lys	Leu	Asn	Val	Thr	Ala	Ile	Asp	Thr
385					390					395				Tyr
														400
Asp	Asp	His	Arg	Met	Ala	Met	Ala	Phe	Ser	Leu	Ala	Ala	Cys	Ala
				405					410					Glu
Val	Pro	Val	Thr	Ile	Arg	Asp	Pro	Gly	Cys	Thr	Arg	Lys	Thr	Phe
			420					425					430	Pro
Asp	Tyr	Phe	Asp	Val	Leu	Ser	Thr	Phe	Val	Lys	Asn			
	435						440							